

1009

OIPE

RAW SEQUENCE LISTING

DATE: 09/28/2001

PATENT APPLICATION: US/09/854,774

TIME: 15:39:11

Input Set : N:\Crf3\RULE60\09854774.txt

Output Set: N:\CRF3\09282001\I854774.raw

2

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:
6 (i) APPLICANT: FROMMER, Wolf-Bernd
8 (ii) TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID
9 TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS
10 CONTAINING A TRANSPORTER AND THEIR USE
12 (iii) NUMBER OF SEQUENCES: 4
14 (iv) CORRESPONDENCE ADDRESS:
15 (A) ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
16 (B) STREET: 1180 Avenue of the Americas
17 (C) CITY: New York
18 (D) STATE: NY
19 (E) COUNTRY: US
20 (F) ZIP: 10036-8403
22 (v) COMPUTER READABLE FORM:
23 (A) MEDIUM TYPE: Floppy disk
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
28 (vi) CURRENT APPLICATION DATA:
C--> 29 (A) APPLICATION NUMBER: US/09/854,774
C--> 30 (B) FILING DATE: 14-May-2001
31 (C) CLASSIFICATION:
33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: 08/964,939
35 (B) FILING DATE: 1998-05-28
37 (A) APPLICATION NUMBER: US 08/007,636
38 (B) FILING DATE: 21-JAN-1993
40 (viii) ATTORNEY/AGENT INFORMATION:
41 (A) NAME: Meilman, Edward A.
42 (B) REGISTRATION NUMBER: 24,735
43 (C) REFERENCE/DOCKET NUMBER: P/951-107
45 (ix) TELECOMMUNICATION INFORMATION:
46 (A) TELEPHONE: (212) 382-0700
47 (B) TELEFAX: (212) 382-0888
48 (C) TELEX: 236925
51 (2) INFORMATION FOR SEQ ID NO: 1:
53 (i) SEQUENCE CHARACTERISTICS:
54 (A) LENGTH: 1685 base pairs
55 (B) TYPE: nucleic acid
56 (C) STRANDEDNESS: single
57 (D) TOPOLOGY: linear
59 (ii) MOLECULE TYPE: cDNA
61 (vi) ORIGINAL SOURCE:
62 (A) ORGANISM: Arabidopsis thaliano
64 (ix) FEATURE:
65 (A) NAME/KEY: CDS

ENTERED

RAW SEQUENCE LISTING

DATE: 09/28/2001

PATENT APPLICATION: US/09/854,774

TIME: 15:39:11

Input Set : N:\Crf3\RULE60\09854774.txt

Output Set: N:\CRF3\09282001\I854774.raw

```

66          (B) LOCATION: 57..1511
67          (D) OTHER INFORMATION: /note= "amino acid transporter"
70          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
72 CTTAAACAT TTATTTTATC TTCTTCTTGT TCTCTCTTTC TCTTTCTCTC ATCACT      56
74 ATG AAG AGT TTC AAC ACA GAA GGA CAC AAC CAC TCC ACG GCG GAA TCC      104
75 Met Lys Ser Phe Asn Thr Glu Gly His Asn His Ser Thr Ala Glu Ser
76   1           5           10           15
78 GGC GAT GCC TAC ACC GTG TCG GAC CCG ACA AAG AAC GTC GAT GAA GAT      152
79 Gly Asp Ala Tyr Thr Val Ser Asp Pro Thr Lys Asn Val Asp Glu Asp
80           20           25           30
82 GGT CGA GAG AAG CGT ACC GGG ACG TGG CTT ACG GCG AGT GCG CAT ATT      200
83 Gly Arg Glu Lys Arg Thr Gly Thr Trp Leu Thr Ala Ser Ala His Ile
84           35           40           45
86 ATC ACG GCG GTG ATA GGC TCC GGA GTG TTG TCT TTA GCA TGG GCT ATA      248
87 Ile Thr Ala Val Ile Gly Ser Gly Val Leu Ser Leu Ala Trp Ala Ile
88           50           55           60
90 GCT CAG CTT GGT TGG ATC GCA GGG ACA TCG ATC TTA CTC ATT TTC TCG      296
91 Ala Gln Leu Gly Trp Ile Ala Gly Thr Ser Ile Leu Leu Ile Phe Ser
92   65           70           75           80
94 TTC ATT ACT TAC TTC ACC TCC ACC ATG CTT GCC GAT TGC TAC CGT GCG      344
95 Phe Ile Thr Tyr Phe Thr Ser Thr Met Leu Ala Asp Cys Tyr Arg Ala
96           85           90           95
98 CCG GAT CCC GTC ACC GGA AAA CGG AAT TAC ACT TAC ATG GAC GTT GTT      392
99 Pro Asp Pro Val Thr Gly Lys Arg Asn Tyr Thr Tyr Met Asp Val Val
100           100           105           110
102 CGA TCT TAC CTC GGT GGT AGG AAA GTG CAG CTC TGT GGA GTG GCA CAA      440
103 Arg Ser Tyr Leu Gly Gly Arg Lys Val Gln Leu Cys Gly Val Ala Gln
104           115           120           125
106 TAT GGG AAT CTG ATT GGG GTC ACT GTT GGT TAC ACC ATC ACT GCT TCT      488
107 Tyr Gly Asn Leu Ile Gly Val Thr Val Gly Tyr Thr Ile Thr Ala Ser
108           130           135           140
110 ATT AGT TTG GTA GCG GTA GGG AAA TCG AAC TGC TTC CAC GAT AAA GGG      536
111 Ile Ser Leu Val Ala Val Gly Lys Ser Asn Cys Phe His Asp Lys Gly
112 145           150           155           160
114 CAC ACT GCG GAT TGT ACT ATA TCG AAT TAT CCG TAT ATG GCG GTT TTT      584
115 His Thr Ala Asp Cys Thr Ile Ser Asn Tyr Pro Tyr Met Ala Val Phe
116           165           170           175
118 GGT ATC ATT CAA GTT ATT CTT AGC CAG ATC CCA AAT TTC CAC AAG CTC      632
119 Gly Ile Ile Gln Val Ile Leu Ser Gln Ile Pro Asn Phe His Lys Leu
120           180           185           190
122 TCT TTT CTT TCC ATT ATG GCC GCA GTC ATG TCC TTT ACT TAT GCA ACT      680
123 Ser Phe Leu Ser Ile Met Ala Ala Val Met Ser Phe Thr Tyr Ala Thr
124           195           200           205
126 ATT GGA ATC GGT CTA GCC ATC GCA ACC GTC GCA GGT GGG AAA GTG GGT      728
127 Ile Gly Ile Gly Leu Ala Ile Ala Thr Val Ala Gly Gly Lys Val Gly
128           210           215           220
130 AAG ACG AGT ATG ACG GGC ACA GCG GTT GGA GTA GAT GTA ACC GCA GCT      776
131 Lys Thr Ser Met Thr Gly Thr Ala Val Gly Val Asp Val Thr Ala Ala
132 225           230           235           240

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/854,774

DATE: 09/28/2001

TIME: 15:39:11

Input Set : N:\Crf3\RULE60\09854774.txt

Output Set: N:\CRF3\09282001\I854774.raw

134	CAA	AAG	ATA	TGG	AGA	TCG	TTT	CAA	GCG	GTT	GGG	GAC	ATA	GCG	TTC	GCC	824
135	Gln	Lys	Ile	Trp	Arg	Ser	Phe	Gln	Ala	Val	Gly	Asp	Ile	Ala	Phe	Ala	
136					245					250						255	
138	TAT	GCT	TAT	GCC	ACG	GTT	CTC	ATC	GAG	ATT	CAG	GAT	ACA	CTA	AGA	TCT	872
139	Tyr	Ala	Tyr	Ala	Thr	Val	Leu	Ile	Glu	Ile	Gln	Asp	Thr	Leu	Arg	Ser	
140				260					265					270			
142	AGC	CCA	GCT	GAG	AAC	AAA	GCC	ATG	AAA	AGA	GCA	AGT	CTT	GTG	GGA	GTA	920
143	Ser	Pro	Ala	Glu	Asn	Lys	Ala	Met	Lys	Arg	Ala	Ser	Leu	Val	Gly	Val	
144				275				280						285			
146	TCA	ACC	ACT	TTT	TTC	TAC	ATC	TTA	TGT	GGA	TGC	ATC	GGC	TAT	GCT		968
147	Ser	Thr	Thr	Thr	Phe	Phe	Tyr	Ile	Leu	Cys	Gly	Cys	Ile	Gly	Tyr	Ala	
148		290					295					300					
150	GCA	TTT	GGA	AAC	AAT	GCC	CCT	GGA	GAT	TTC	CTC	ACA	GAT	TTC	GGG	TTT	1016
151	Ala	Phe	Gly	Asn	Asn	Ala	Pro	Gly	Asp	Phe	Leu	Thr	Asp	Phe	Gly	Phe	
152	305					310					315					320	
154	TTC	GAG	CCC	TTT	TGG	CTC	ATT	GAC	TTT	GCA	AAC	GCT	TGC	ATC	GCT	GTC	1064
155	Phe	Glu	Pro	Phe	Trp	Leu	Ile	Asp	Phe	Ala	Asn	Ala	Cys	Ile	Ala	Val	
156				325					330					335			
158	CAC	CTT	ATT	GGT	GCC	TAT	CAG	GTG	TTC	GCG	CAG	CCG	ATA	TTC	CAG	TTT	1112
159	His	Leu	Ile	Gly	Ala	Tyr	Gln	Val	Phe	Ala	Gln	Pro	Ile	Phe	Gln	Phe	
160				340					345					350			
162	GTT	GAG	AAA	AAA	TGC	AAC	AGA	AAC	TAT	CCA	GAC	AAC	AAG	TTC	ATC	ACT	1160
163	Val	Glu	Lys	Lys	Cys	Asn	Arg	Asn	Tyr	Pro	Asp	Asn	Lys	Phe	Ile	Thr	
164			355					360					365				
166	TCT	GAA	TAT	TCA	GTA	AAC	GTA	CCT	TTC	CTT	GGA	AAA	TTC	AAC	ATT	AGC	1208
167	Ser	Glu	Tyr	Ser	Val	Asn	Val	Pro	Phe	Leu	Gly	Lys	Phe	Asn	Ile	Ser	
168		370					375					380					
170	CTC	TTC	AGA	TTG	GTG	TGG	AGG	ACA	GCT	TAT	GTG	GTT	ATA	ACC	ACT	GTT	1256
171	Leu	Phe	Arg	Leu	Val	Trp	Arg	Thr	Ala	Tyr	Val	Val	Ile	Thr	Thr	Val	
172	385				390						395					400	
174	GTA	GCT	ATG	ATA	TTC	CCT	TTC	TTC	AAC	GCG	ATC	TTA	GGT	CTT	ATC	GGA	1304
175	Val	Ala	Met	Ile	Phe	Pro	Phe	Phe	Asn	Ala	Ile	Leu	Gly	Leu	Ile	Gly	
176				405					410					415			
178	GCA	GCT	TCC	TTC	TGG	CCT	TTA	ACG	GTT	TAT	TTC	CCT	GTG	GAG	ATG	CAC	1352
179	Ala	Ala	Ser	Phe	Trp	Pro	Leu	Thr	Val	Tyr	Phe	Pro	Val	Glu	Met	His	
180				420					425					430			
182	ATT	GCA	CAA	ACC	AAG	ATT	AAG	AAG	TAC	TCT	GCT	AGA	TGG	ATT	GCG	CTG	1400
183	Ile	Ala	Gln	Thr	Lys	Ile	Lys	Lys	Tyr	Ser	Ala	Arg	Trp	Ile	Ala	Leu	
184				435				440						445			
186	AAA	ACG	ATG	TGC	TAT	GTT	TGC	TTG	ATC	GTC	TCG	CTC	TTA	GCT	GCA	GCC	1448
187	Lys	Thr	Met	Cys	Tyr	Val	Cys	Leu	Ile	Val	Ser	Leu	Leu	Ala	Ala	Ala	
188		450					455					460					
190	GGA	TCC	ATC	GCA	GGA	CTT	ATA	AGT	AGT	GTC	AAA	ACC	TAC	AAG	CCC	TTC	1496
191	Gly	Ser	Ile	Ala	Gly	Leu	Ile	Ser	Ser	Val	Lys	Thr	Tyr	Lys	Pro	Phe	
192	465					470					475					480	
194	CGG	ACT	ATG	CAT	GAG	TGAGTTTGAG	ATCCTCAAGA	GAGTCAAAAA	TATATGTAGT								1551
195	Arg	Thr	Met	His	Glu												
196					485												
198	AGTTTGGTCT	TTCTGTTAAA	CTATCTGGTG	TCTAAATCCA	ATGAGAATGC	TTTATTGCTA											1611

RAW SEQUENCE LISTING

DATE: 09/28/2001

PATENT APPLICATION: US/09/854,774

TIME: 15:39:11

Input Set : N:\Crif3\RULE60\09854774.txt

Output Set: N:\CRF3\09282001\I854774.raw

```

200 AAAC TTCATG AATCTCTCTG TATCTACATC TTTCAATCTA ATACATATGA GCTCTTCCAA 1671
202 AAAAAAAAAA AAAA 1685
205 (2) INFORMATION FOR SEQ ID NO: 2:
207     (i) SEQUENCE CHARACTERISTICS:
208         (A) LENGTH: 485 amino acids
209         (B) TYPE: amino acid
210         (D) TOPOLOGY: linear
212     (ii) MOLECULE TYPE: protein
214     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
216 Met Lys Ser Phe Asn Thr Glu Gly His Asn His Ser Thr Ala Glu Ser
217   1      5      10      15
219 Gly Asp Ala Tyr Thr Val Ser Asp Pro Thr Lys Asn Val Asp Glu Asp
220      20      25      30
222 Gly Arg Glu Lys Arg Thr Gly Thr Trp Leu Thr Ala Ser Ala His Ile
223      35      40      45
225 Ile Thr Ala Val Ile Gly Ser Gly Val Leu Ser Leu Ala Trp Ala Ile
226      50      55      60
228 Ala Gln Leu Gly Trp Ile Ala Gly Thr Ser Ile Leu Leu Ile Phe Ser
229      65      70      75      80
231 Phe Ile Thr Tyr Phe Thr Ser Thr Met Leu Ala Asp Cys Tyr Arg Ala
232      85      90      95
234 Pro Asp Pro Val Thr Gly Lys Arg Asn Tyr Thr Tyr Met Asp Val Val
235      100     105     110
237 Arg Ser Tyr Leu Gly Gly Arg Lys Val Gln Leu Cys Gly Val Ala Gln
238      115     120     125
240 Tyr Gly Asn Leu Ile Gly Val Thr Val Gly Tyr Thr Ile Thr Ala Ser
241      130     135     140
243 Ile Ser Leu Val Ala Val Gly Lys Ser Asn Cys Phe His Asp Lys Gly
244 145      150     155     160
246 His Thr Ala Asp Cys Thr Ile Ser Asn Tyr Pro Tyr Met Ala Val Phe
247      165     170     175
249 Gly Ile Ile Gln Val Ile Leu Ser Gln Ile Pro Asn Phe His Lys Leu
250      180     185     190
252 Ser Phe Leu Ser Ile Met Ala Ala Val Met Ser Phe Thr Tyr Ala Thr
253      195     200     205
255 Ile Gly Ile Gly Leu Ala Ile Ala Thr Val Ala Gly Gly Lys Val Gly
256      210     215     220
258 Lys Thr Ser Met Thr Gly Thr Ala Val Gly Val Asp Val Thr Ala Ala
259 225      230     235     240
261 Gln Lys Ile Trp Arg Ser Phe Gln Ala Val Gly Asp Ile Ala Phe Ala
262      245     250     255
264 Tyr Ala Tyr Ala Thr Val Leu Ile Glu Ile Gln Asp Thr Leu Arg Ser
265      260     265     270
267 Ser Pro Ala Glu Asn Lys Ala Met Lys Arg Ala Ser Leu Val Gly Val
268      275     280     285
270 Ser Thr Thr Thr Phe Phe Tyr Ile Leu Cys Gly Cys Ile Gly Tyr Ala
271      290     295     300
273 Ala Phe Gly Asn Asn Ala Pro Gly Asp Phe Leu Thr Asp Phe Gly Phe
274 305      310     315     320

```

RAW SEQUENCE LISTING

DATE: 09/28/2001

PATENT APPLICATION: US/09/854,774

TIME: 15:39:11

Input Set : N:\Crf3\RULE60\09854774.txt

Output Set: N:\CRF3\09282001\I854774.raw

```

276 Phe Glu Pro Phe Trp Leu Ile Asp Phe Ala Asn Ala Cys Ile Ala Val
277           325           330           335
279 His Leu Ile Gly Ala Tyr Gln Val Phe Ala Gln Pro Ile Phe Gln Phe
280           340           345           350
282 Val Glu Lys Lys Cys Asn Arg Asn Tyr Pro Asp Asn Lys Phe Ile Thr
283           355           360           365
285 Ser Glu Tyr Ser Val Asn Val Pro Phe Leu Gly Lys Phe Asn Ile Ser
286           370           375           380
288 Leu Phe Arg Leu Val Trp Arg Thr Ala Tyr Val Val Ile Thr Thr Val
289 385           390           395           400
291 Val Ala Met Ile Phe Pro Phe Phe Asn Ala Ile Leu Gly Leu Ile Gly
292           405           410           415
294 Ala Ala Ser Phe Trp Pro Leu Thr Val Tyr Phe Pro Val Glu Met His
295           420           425           430
297 Ile Ala Gln Thr Lys Ile Lys Lys Tyr Ser Ala Arg Trp Ile Ala Leu
298           435           440           445
300 Lys Thr Met Cys Tyr Val Cys Leu Ile Val Ser Leu Leu Ala Ala Ala
301           450           455           460
303 Gly Ser Ile Ala Gly Leu Ile Ser Ser Val Lys Thr Tyr Lys Pro Phe
304 465           470           475           480
306 Arg Thr Met His Glu
307           485

```

309 (2) INFORMATION FOR SEQ ID NO: 3:

311 (i) SEQUENCE CHARACTERISTICS:

312 (A) LENGTH: 1740 base pairs

313 (B) TYPE: nucleic acid

314 (C) STRANDEDNESS: single

315 (D) TOPOLOGY: linear

317 (ii) MOLECULE TYPE: cDNA

319 (vi) ORIGINAL SOURCE:

320 (A) ORGANISM: Arabidopsis thaliana

322 (ix) FEATURE:

323 (A) NAME/KEY: CDS

324 (B) LOCATION: 80..1558

325 (D) OTHER INFORMATION: /product= "amino acid transporter"

328 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

330 CTATTTTATA ATTCCTCTTC TTTTGTTCA TAGCTTGTA ATTATAGTCT TATTTCTCTT      60
332 TAAGGCTCAA TAAGAGGAG ATG GGT GAA ACC GCT GCC GCC AAT AAC CAC CGT      112
333           Met Gly Glu Thr Ala Ala Ala Asn Asn His Arg
334           1           5           10
336 CAC CAC CAC CAT CAC GGC CAC CAG GTC TTT GAC GTG GCC AGC CAC GAT      160
337 His His His His His Gly His Gln Val Phe Asp Val Ala Ser His Asp
338           15           20           25
340 TTC GTC CCT CCA CAA CCG GCT TTT AAA TGC TTC GAT GAT GAT GGC CGC      208
341 Phe Val Pro Pro Gln Pro Ala Phe Lys Cys Phe Asp Asp Asp Gly Arg
342           30           35           40
344 CTC AAA AGA ACT GGG ACT GTT TGG ACC GCG AGC GCT CAT ATA ATA ACT      256
345 Leu Lys Arg Thr Gly Thr Val Trp Thr Ala Ser Ala His Ile Ile Thr
346           45           50           55

```

VERIFICATION SUMMARY

DATE: 09/28/2001

PATENT APPLICATION: US/09/854,774

TIME: 15:39:12

Input Set : N:\Crf3\RULE60\09854774.txt

Output Set: N:\CRF3\09282001\I854774.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]